

Automated robot-assisted wide-field optical coherence tomography using structured light camera

XIAOCHEN LI, YONG HUANG,*  AND QUN HAO

School of Optics and Photonics, Beijing Institute of Technology, No.5 South Zhongguancun Street, Haidian, Beijing, 100081, China

*huangyong2015@bit.edu.cn

Abstract: Optical coherence tomography (OCT) is a promising real-time and non-invasive imaging technology widely utilized in biomedical and material inspection domains. However, limited field of view (FOV) in conventional OCT systems hampers their broader applicability. Here, we propose an automated system integrating a structured light camera and robotic arm for large-area OCT scanning. The system precisely detects tissue contours, automates scan path generation, and enables accurate scanning of expansive sample areas. The proposed system consists of a robotic arm, a three-dimensional (3D) structured light camera, and a customized portable OCT probe. The 3D structured light camera is employed to generate a precise 3D point cloud of the sample surface, enabling automatic planning of the scanning path for the robotic arm. Meanwhile, the OCT probe is mounted on the robotic arm, facilitating scanning of the sample along the predetermined path. Continuous OCT B-scans are acquired during the scanning process, facilitating the generation of high-resolution and large-area 3D OCT reconstructions of the sample. We conducted position error tests and presented examples of 3D macroscopic imaging of different samples, such as *ex vivo* kidney, skin and leaf blade. The robotic arm can accurately reach the planned positions with an average absolute error of approximately 0.16 mm. The findings demonstrate that the proposed system enables the acquisition of 3D OCT images covering an area exceeding 20 cm², indicating wide-ranging potential for utilization in diverse domains such as biomedical, industrial, and agricultural fields.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Optical coherence tomography (OCT) is a non-contact and non-invasive imaging technology. It provides real-time, high-resolution cross-sectional images, and volumetric reconstructions with a penetration depth of several millimeters in biological tissues or materials [1–4]. The primary applications of OCT are in biomedical imaging and diagnosis, including fields such as ophthalmology [5], dermatology [6], oncology [7], and dentistry [8]. Furthermore, OCT has been utilized in non-biomedical areas such as art conservation for non-destructive analysis [9], botany [10], fruit quality assessment [11].

However, conventional OCT suffers from a limited imaging field of view (FOV), typically restricted to a few centimeters [2,12,13]. This limitation confines OCT to imaging small areas, making it challenging to capture the structures of interest in a single acquisition. In many biomedical applications, there is a demand for high-depth-resolution imaging with a wide and deep FOV to assess large biological samples, such as whole eye assessment [14], skin [15], and whole brain vascular visualization [16,17]. Likewise, in non-biomedical fields, the non-destructive and accurate visualization of large-area samples is also imperative [18]. Therefore, the expansion of the FOV in OCT has become an increasingly important research objective.

To achieve a wide FOV in OCT imaging, a promising strategy involves the integration of a long coherence length scanning source with a wide lateral FOV scanning objective [19,20]. This

combination allows for the acquisition of comprehensive OCT 3D macroscopic images covering wide areas in a single scan. However, it should be noted that optimizing the system for a long Rayleigh range often necessitates a trade-off in lateral resolution, resulting in a reduction in the effective resolution of the scattered sample. Therefore, striking a balance between a large lateral FOV and high resolution is of paramount importance in clinical research.

An alternative strategy is to employ the stitching of multiple high-resolution OCT volumes for the reconstruction of large FOV images [21]. Essentially, the effective FOV improvement is enabled by the lateral FOV extension while the axial FOV of OCT image from every single volume is the same. This technique usually entails the use of XYZ linear transformation stages to finely adjust the probe's motion [22]. However, it should be noted that these systems often exhibit limited degrees of freedom, thereby constraining scanning flexibility and impeding workspace adaptability.

Recently, there has been notable progress in the development of miniaturized handheld devices, enabling operators to perform scanning over a wider range with increased flexibility [23–25]. However, the presence of inherent hand tremors among operators introduces unintended distance variations between the OCT probe and the tissue surface, leading to undesirable consequences such as image blurring, the emergence of artifacts, and even signal loss [26,27]. Furthermore, the lack of precise spatial localization information for the OCT probe poses a significant challenge, hindering the accurate determination of the corresponding position in space for acquired images. As a result, this limitation poses obstacles to subsequent tasks such as large-scale volumetric reconstruction and the realization of comprehensive 3D visualization.

To address the challenges outlined above, various robotic arm assisted imaging systems not limited to OCT imaging were proposed. Our previous work utilized a 6 DoF robotic arm to drive a compact OCT imaging probe through eye-in-hand configuration [28]. Similarly, He et al. utilized a robotic arm to facilitate optical coherence tomography angiography imaging of microvasculature in the entire brain of mice [17]. Göb et al. combined a robotic arm with a custom-built high speed 3.3 MHz-OCT system to create a robot-assisted OCT setup with the aim of covering larger areas [29]. The successful practice of these works demonstrates the potential of expanding the imaging FOV with robotic arms.

However, previous work [17,28,29] on robotic arm assisted OCT imaging requires manual positioning of the OCT probe. The reason is that they did not leverage other imaging modalities to accurately extract target sample surface contours, and lacked the capability to pre-determine scan paths. Discontinuous scanning was performed, wherein OCT C-scan was acquired at each individual scanning point before moving to the next. Due to the limited FOV of one individual OCT imaging volume, an OCT volume pre-scan was conducted at each acquisition point to optimize the probe alignment to sample using an object surface detection algorithm based on OCT images. However, challenges arise due to strong reflections at the top of the OCT images and the intricate contours of the tissue surface, often leading to inaccurate tissue surface detection, may result in errors in robot positioning and the loss of OCT image signal. In addition, post 3D reconstruction requires additional 3D image registration between individual volumes.

Autonomous scanning has always been desirable. For example, Zhang et al. [30] used the da Vinci robot to perform autonomous large-area pCLE scans over a user-defined area through eye-out-of-hand probe position calibration and image registration. Ortiz et al. [31] focused on retinal OCT image stitching and successfully expanded the FOV of the system by exploiting the robot's rotational capabilities and gaze tracking control through eye-in-hand configuration and image registration between acquired volumes. Ma [32] utilized an RGB depth camera and the OCT images working together in hybrid to generate the autonomous scan path. First, the RGB depth camera was used to generate a coarse scan path within a plane. Then OCT images were used in closed-loop to follow the sample surface during the scanning process, resulting an effective scan path that can follow sample surface contour. However, as mentioned in [32]

since the eventual scan path relies on OCT image height compensation, variations of OCT image quality under different scenarios will cause height detection inaccuracy and sometimes even erroneous detection, resulting a non-smooth and non-uniform scan path that will cause image distortion and even signal loss.

In this study, we continued the idea of stitching to extend the lateral FOV and made advancements to the robotic-arm-assisted OCT system by integrating a high-precision structured light camera. This integration allows for accurate perception of the sample's surface topology, automated acquisition and segmentation of contour point clouds, and generation of scanning paths. This development eliminates the need for manual alignment of the probe to the target area, as operators can now select any region of interest within the structured light camera's FOV for scanning. By combining machine vision technology with robotic arm control, we have improved the autonomy and intelligence of the robotic arm in complex working environments and tasks, thereby enhancing its adaptability and practicality. What's more, allowing the operator to directly see the planned movement path of the robotic arm before scanning ensures safe scanning without collision of the probe with the sample. As a result, compared to previous work, our work provides safe and accurate method for wide-field OCT imaging.

To evaluate the performance of the proposed system, we conducted demonstrations of 3D macroscopic imaging on various samples including *in vivo* skin, *ex vivo* kidney, as well as leaf blade, citrus, and ceramics. In our work, the scan time of our system for an area of approximate 19.4 cm^2 ($81.5\text{ mm} \times 23.8\text{ mm}$) consisting of 15000 B-mode images (4.2 mm lateral FOV with 0.2 mm redundancy, 512×512 pixels) is about 8 mins, which depends on the sampling intervals determined by the OCT imaging speed (31 fps) and probe moving speed (1 mm/s). Theoretically, the largest imaging lateral FOV is $248\text{ mm} \times 164\text{ mm}$, which is limited by the FOV of structured light camera. The proposed system is expected to expand the applications of OCT imaging in the biomedical field and beyond.

2. Materials and methods

2.1. System overview

As shown in Fig. 1(a), OCT images were acquired using a homebuilt spectral domain OCT (SD-OCT) system with a broadband super luminescent diode (SLD) (S5FC1021S, Thorlabs, Inc.) with a central wavelength of 1310 nm and an output power of $\sim 12.5\text{ mW}$ and -3 dB bandwidth of 85 nm. In the fiber-based OCT system, a 50/50 fiber coupler (BXC25, Thorlabs, Inc.) is used to separate the light to the sample arm and the reference arm. The reflected light was transmitted to a high-speed spectrometer (Cobra 1300, Wasatch Photonics, Inc.). The spectrometer detected the spectral range of 1175 ~ 1420 nm with a maximum speed of 76 kHz. The system was configured to acquire A-lines at a rate of 32 kHz, with each B-scan consisting of 1024 A-lines with a depth of 1024 pixels. The OCT system has a lateral and axial resolution of $\sim 24.8\text{ }\mu\text{m}$ and $\sim 12.5\text{ }\mu\text{m}$ in air, respectively.

In the customized OCT probe, the beam from the optical fiber is collimated by a collimator and then directed to a two-dimensional micro-electromechanical systems (MEMS) scanning mirror (A7B1.1, Mirrorcle Tech., Inc.) on the front focal plane of the scanning lens combination for telecentric scanning. The beam was continuously deflected by the MEMS scanning mirror and moved over the sample surface to acquire slices (B-scan) covering a lateral FOV of 4 mm with 1024×1024 pixels. B-scan image is down sampled to 512×512 pixels to save. The sample was positioned under the scanning head and the robotic arm was moved along the sample surface with a working distance of 30 mm between the surface and the scanning head lens. A long-pass dichroic mirror (DMR-950LP, LBTEK, Inc.) with a truncated wavelength of 950 nm was used to separate the OCT optical path from the RGB camera optical path. In addition, the RGB camera integrated into the scan head optical path covers the scanning FOV, providing a real-time view of the imaging area. The enclosure of the OCT probe is 3D printed from black resin and the

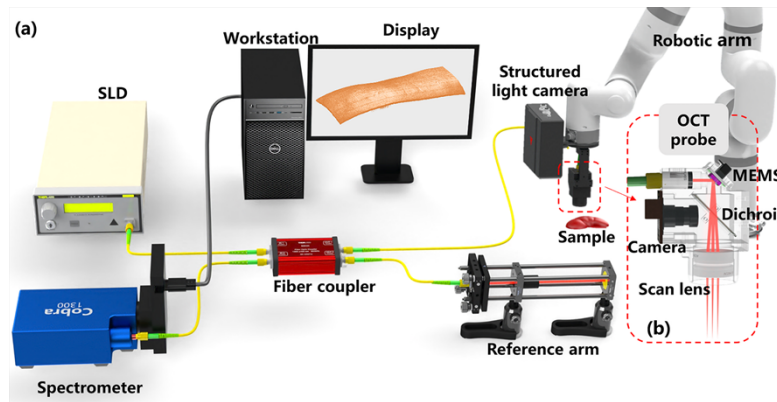


Fig. 1. Schematic of the system setup. (a) The components include a super luminescent diode (SLD), a 50/50 fiber optic coupler, a high-resolution spectrometer, a reference arm, a custom-designed OCT scanning probe with integrated camera, a structured light camera, a powerful workstation for data processing and analysis, a display and a precise 6-DoF robotic arm for controlled probe movement. (b) The schematic diagram presents the detailed optical configuration of the custom OCT scanning probe, illustrating the arrangement of lenses, MEMS, and other optical elements.

probe weighs 150 grams, is 53 mm long, 34 mm wide and 72 mm high. We implemented MEMS scanning mirror to reduce the probe size, and employed 3D-printed resin materials for the probe frame to achieve weight reduction.

The 3D structured light camera (RVC-I2370, RVBUST, Inc.) utilizes 465 nm blue stripe structured light and an advanced point cloud synthesis algorithm to capture images of a surface. It offers a high single point repeatability of 0.04 mm and a FOV of $248 \times 164 \text{ mm}^2$ at a distance of 400 mm. The camera captures RGB images and depth images at a resolution of 1440×1080 . The OCT probe and structured light camera are mounted on a 6-DoF robotic arm (xArm7, UFACTORY, Inc.) with a payload capacity of 3.5 kg and a repeatability of $\pm 0.1 \text{ mm}$. The system's software is developed in C++ with a graphical user interface (GUI) implemented in QT, enabling OCT acquisition and processing, 3D structured light camera control, point cloud generation, and robotic arm control. The optical equipment, controllers, and computer components are integrated into a customized medical device cart, designed for easy use in clinical environments.

2.2. System calibrations

In large-field OCT scanning, a high-precision structured light camera (eye) is used to sense the surface topography of the sample and to navigate the robotic arm (hand). System calibration is a necessary step to ensure “hand-eye coordination” and system accuracy. As shown in Fig. 2(a), there are several coordinate systems: the base coordinate system of the robotic arm (denoted by R), the coordinate system with the center point of the flange at the end of the robotic arm as the origin (denoted by E), the structured light camera coordinate system fixed on top of the robotic arm (denoted by C), the calibrator coordinate system (denoted by H) and the OCT coordinate system (denoted by O). By hand-eye calibration between the camera coordinate system (C) and the robot end coordinate system (E), we can obtain the coordinates of the scan point in robot space. The position translation between the OCT imaging beam focus and the robot flange is achieved by calibrating the OCT coordinates (O) to the robot end coordinate system (E), so that the OCT probe is correctly aligned with the pre-planned scan point, as described in detail in our previous work [28].

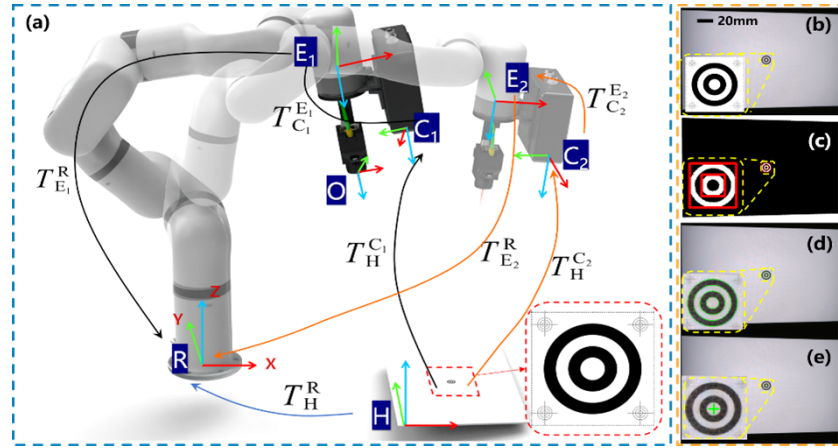


Fig. 2. System calibration schematic. (a) Eye-in-hand configuration. The structured light camera generates and projects specific stripes into the scene and captures images at different locations for subsequent analysis. (b)-(e) The process involves detecting and precisely locating the center of a specific marker circle on the calibration plate, and the circle center coordinates are used to estimate the camera's pose relative to the plate. The pose of the camera at multiple locations is then used to solve a hand-eye transformation matrix that relates the camera coordinates to the robot end-effector coordinates.

Prior to obtaining the point cloud from the structured light camera, it is necessary to perform registration and fusion of the depth and color images, converting pixel coordinates into real world coordinates. This process is achieved through intrinsic calibration of the structured light camera. The camera undergoes intrinsic calibration prior to shipment and does not require any further calibration.

In this case, we need to find the coordinate transformation relationship between the camera coordinate system and the end-of-arm coordinate T^E_C . The structured light camera is fixed at the end of the robotic arm and moves along with the arm, keeping their positions relatively constant. We use the eye-in-hand calibration method.

First, we place a specific calibration pattern in the FOV of the camera, which is fixed in position. The robotic arm moves the camera in different poses and positions to capture specific calibration patterns. The pixel coordinates of the center point of the calibration pattern are also obtained.

For any two positions, the coordinates of the calibration pattern in the R coordinate system and the H coordinate system remain unchanged, as the calibration pattern and the base of the robotic arms remain stationary. However, the coordinates of the calibration pattern in the E coordinate system and the C coordinate system change as the robotic arms move. Furthermore, during the movement of the robotic arms, the relative position between the camera coordinate system C and the end effector coordinate system E remains fixed. Therefore, the following equation is available:

$$X = T^{E_1}_{C_1} = T^{E_2}_{C_2} \quad (1)$$

$$T^R_H = T^R_{E_1} \times X \times T^{C_1}_H = T^R_{E_2} \times X \times T^{C_2}_H \quad (2)$$

$$(T^{C_2}_H)^{-1} \times T^R_{E_2} \times X = X \times T^{C_1}_H \times (T^{C_1}_H)^{-1} \quad (3)$$

The coordinate transformation from the end coordinate system of the robotic arm to the base coordinate system of the robotic arm can be obtained by reading the pose parameters of the robotic arm and the robot kinematic model. The conversion of the calibration pattern coordinate to the camera coordinate T^C_H can be obtained from the camera intrinsic matrix and the acquired

pictures. Let $A = (T_{E_2}^R)^{-1} \times T_{E_1}^R$, $B = T_H^{C_2} \times (T_H^{C_1})^{-1}$, so that the problem is formulated as a linear equation $AX = XB$. This equation is solved by moving the camera to different positions (we set 30 positions) to obtain multiple sets of equations to obtain X [33,34].

The structured light camera captures images of the calibration plate at different locations to identify a specific marker circle and obtain its central point coordinates in the camera coordinate system. OpenCV library is utilized for circular marker detection within the robot workspace. The captured color image of the marker circle (Fig. 2(b)) undergoes preprocessing steps such as noise removal, grayscale conversion, and binary image generation using the OTSU method [35]. The binary image is then processed through connected component extraction and sorting based on pixel count and aspect ratio, resulting in a rough localization of the marker circle's rectangular range (Fig. 2(c)). Within each rectangular range, an improved canny method is applied to extract subpixel edge contours [36]. Closed contours meeting specific criteria are identified (Fig. 2(d)), and their points are fitted with an ellipse to obtain the ellipse's center coordinates (Fig. 2(e)). The obtained pixel coordinates of the circle's center are converted to coordinates in the structured light camera coordinate system using the internally calibrated reference matrix provided by the camera manufacturer.

2.3. Scan area selection and scan path planning

The presented system adopts an eye-in-hand installation method. Before scanning, the robotic arm with the fixed structured light camera automatically moves to the initial position (Fig. 3(a)), and the sample is placed in the center area of the structured light camera's FOV. The structured light camera projects blue stripes onto the sample surface, captures the sample's color and depth images, and generates a point cloud of the sample surface. On the GUI containing the structured light camera's color view, the operator can use the mouse to select the region of interest (ROI) (green box).

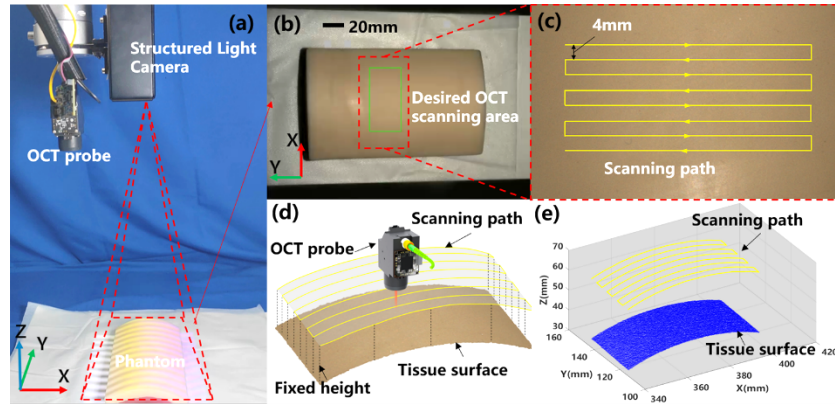


Fig. 3. Scan area selection and scan path planning. (a) The structured light camera projects a stripe pattern into the scene and captures the image. (b) The user marks the scan area with a green box on the GUI interface. (c) The scan path and direction are generated automatically. The spacing of the two scan lines is 4 mm. (d) The structured light camera generates a color point cloud of the sample surface. During the scanning process, the OCT probe should be at a fixed distance from the surface. (e) The spatial position of the sample surface (blue) and the generated scan path (yellow) are visualized, depicting the relationship between the scanned area and the planned path.

The system automatically plans a scanning path to ensure full coverage of the ROI (Fig. 3(c)). The OCT probe performs raster scanning along the direction indicated by the scanning line arrow. Currently, the scanner was positioned downwards matching the normal of sample resting table.

During the scanning process, the probe only moves with XYZ translational movement while the orientation of the scanner is fixed instead of dynamically adjusted to match the normal of the sample surface. The OCT B-scan has a FOV of 4.2 mm, and the distance between adjacent scanning lines is set to 4 mm to ensure overlap between two adjacent scanning lines. The number and length of scanning lines are limited by the scanning frame to ensure that they do not exceed its boundaries.

Figure 3(d) illustrates the surface contour point cloud of the ROI. To prevent sample loss or contact with the OCT probe during scanning, a fixed distance of 30 mm is maintained between the OCT probe and the sample surface. The blue points represent the surface contour of the sample (only retaining the position information), and the yellow path points above indicate the automatically computed scanning path points (Fig. 3(e)).

2.4. 3-D reconstruction and segmentation

After the scan path is planned, the robotic arm needs to move accurately along the planned path and acquire OCT images. Since the coordinates of the generated scan path points are in the coordinate system of the structured light camera, the coordinates are first converted to the coordinate system of the robotic arm by the transformation matrix from the camera to the robotic arm obtained through system calibration. Then, the robotic arm moves the OCT probe along the scan path at a constant speed of 1 mm/s. The moving speed can be set by the user. If the speed is too fast, the interval between OCT images increases, the sampling rate decreases, and the slow-moving speed reduces the efficiency of the system. Based on experience, we chose a moving speed of 1 mm/s, which ensures an acceptable sampling rate (the interval between two adjacent B-scan is $32\ \mu\text{m}$) and efficiency. As shown in Fig. 4(a), during the movement of the probe, continuous two-dimensional OCT images (OCT B-scans) are acquired, and the OCT probe always maintains the optimal distance from the tissue surface. After the scanning process, 3D OCT image reconstruction is performed using the acquired 2D OCT images with associated position information. Specifically, each OCT B-scan is correlated with the corresponding position of the image acquisition. By combining the pixel positions of OCT images with OCT probe position parameters, the real-world position of each pixel can be calculated, and the pixel points can be mapped to the actual space for 3D visualization of the tissue structure.

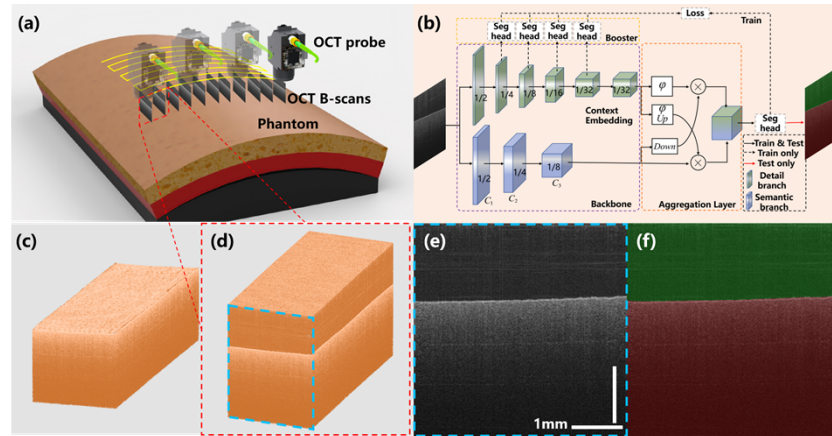


Fig. 4. Schematic diagram of image segmentation and 3D reconstruction based on the BiSeNetV2 network. (a) The OCT probe moves uniformly along the predetermined path, capturing consecutive OCT B-scans. (b) Architecture of the BiSeNetV2 model. (c) 3D visualization after background removal through segmentation. (d) 3D visualization prior to segmentation. (e) OCT B-scan image. (f) Segmentation prediction result.

Before performing 3D reconstruction, we need to segment the OCT B-scans to remove the background above the sample surface, in order to better visualize the 3D reconstruction. In consideration of the need for real-time data processing and limitations of computational resources, a lightweight BiSeNetV2 model [37] was selected for simple semantic segmentation. Its overall architecture is illustrated in Fig. 4(b). In the backbone of the network, there are two branches: one is designed to capture spatial details with a wide channel and shallow layers, called the detail branch. In contrast, the other branch is used to extract classification semantics with narrow channels and deep layers, called the semantic branch.

Figure 4(e) and 4(f) show an OCT B-scan and its corresponding segmentation result, where the green region represents the background and the red region represents the sample. It can be observed that the lightweight semantic segmentation network can accurately segment the OCT images. Figure 4(d) displays a local area of the 3D reconstruction, where the sample surface is wrapped and occluded by the background before segmentation. Figure 4(c) presents the 3D reconstruction after removing the background, which reveals the previously occluded sample surface.

3. Experiments and Results

3.1. Evaluation of the robotic positioning accuracy

After the hand-eye calibration process, the transformation relationship between the robot end effector coordinate system (E) and the structured light camera coordinate system (C) was obtained. The ability of the robotic arm to accurately follow the scanning path generated by the structured light camera will directly affect the final OCT 3D reconstruction result. In our previous work, we tested the repeatable positioning accuracy of the robotic arm at 0.112 mm, which is close to the official figure. In order to directly evaluate the accuracy of the robotic arm moving the probe to align any target point in space under the navigation of the structured light camera, we performed absolute positioning accuracy tests. First, the alignment of the center of the OCT scanning beam with the center of the camera FOV is achieved through precise mechanical design and adjustment. A plate with a specific marker circle is placed in the FOV of the structured light camera. The structured light camera automatically identifies and locates the center of the marked circle in the structured light camera image, and then guides the robotic arm to move the OCT probe to align with the center of the marked circle (Fig. 5(a)).

Figure 5(b) illustrates the image obtained from the built-in RGB camera when the OCT probe is aligned with the marked circle. The green cross represents the camera's FOV center, and the horizontal error (XY) is calculated based on the distance between its center and the center of the marker circle (yellow point). The axial error (Z) is determined by marking the edge position of the marked circle on the OCT image. The ideal height (H) is defined as 240 pixels from the top of the image, and the difference between H and the distance (h) from the center of the upper edge of the marked circle to the top of the image represents the axial error. We conducted 30 repeated experiments with the calibration plate placed in various positions. Figure 5(d)-(f) display scatter plots representing the error data in the XY, XZ, and YZ planes. These scatter points exhibit a symmetric distribution around the origin (0,0), and the enclosing circle has a small radius of approximately 0.16 mm.

3.2. Phantom experiment

To assess the accuracy of its image-guided large FOV automated scanning, we conducted phantom experiments. A curved skin phantom with a black paper strip (size 10 mm × 50 mm) attached to it was used as the test object, as depicted in Fig. 6(a). The planned scanning lines covered an area of 81.5 mm × 23.8 mm. A total of about 15,000 B-scans were collected during the whole scanning process, which took about 8 minutes. We employ C++ and the PCL (Point Cloud Library) [40] to

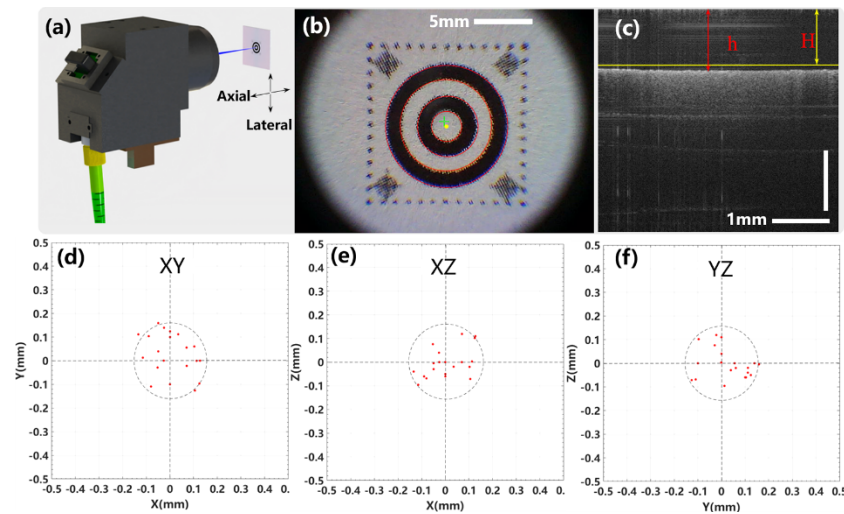


Fig. 5. Evaluation of the robotic arm's positioning accuracy was performed. (a) The OCT probe, equipped with an RGB camera, was aligned with a specific marker circle under the guidance of the structured light camera. (b) The captured image by the RGB camera displays the marker circle, with the green cross indicating the center of the camera's FOV, and the yellow dot representing the center of the marker circle. (c) The depth error was estimated using the OCT image of the marker board. Here, H denotes the predetermined ideal height, while h corresponds to the actual height measured from the top edge of the marker board to the top of the OCT image. (d)-(f) The error data was projected onto the XY, XZ, and YZ planes, respectively, providing an intuitive visualization of the positioning accuracy.

project OCT image pixels onto their corresponding spatial positions. Typically, the reconstruction process takes around 2.5 minutes. The OCT image pixels were extracted and converted to real-space point cloud data, which was visualized in 3D using CloudCompare software (Fig. 6(b)). Measurements of the black paper strip's length and width were performed using measurement tool, yielding a width of 9.98 mm and a length of 49.88 mm. These measurements closely matched the actual size of the strip. Figure 6(c) shows a comparison between the surface point cloud obtained by the structured light camera (blue) and the surface point cloud obtained by OCT (red), and we calculated the Hausdorff distance (the maximum neighboring distance between two point clouds) to be 0.48 mm, demonstrating the accuracy of the OCT reconstruction.

The stitched OCT *en face* image of the phantom's surface was presented in Fig. 6(d). Additionally, a depth-encoded map was generated based on the actual coordinates of surface points in real space, revealing the surface geometry of the phantom in the real-world environment (Fig. 6(e)). [Visualization 1](#) records the real-time B-Scans during the scanning process. Exemplary OCT depth scans (2D cross-sections) Fig. 6(f) and Fig. 6g with high resolution image shown in [Dataset 1](#) (Ref. [43]) were provided to demonstrate the overall OCT quality.

3.3. Biomedical applications

OCT holds great promise in renal transplantation, enabling high-resolution imaging of renal structure and function. This is crucial for pre- and post-transplant evaluation of the transplanted kidney and prediction of its long-term survival [38]. Wide-field OCT imaging in *ex vivo* renal imaging offers significant advantages and valuable insights. In our study, an *ex vivo* porcine kidney was selected, and a specific region (81.1 mm × 24.1 mm) was scanned (Fig. 7(a)). Continuous OCT B-scans were collected (Fig. 7(f)) and processed using semantic segmentation to remove

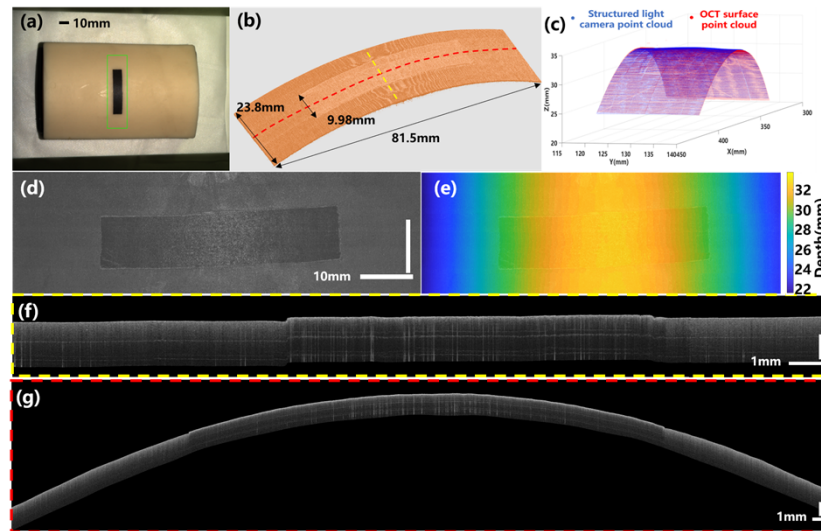


Fig. 6. Experimental results of scanning skin phantom. (a) Photograph of the skin phantom with black paper strips pasted on it as a size reference. (b) 3D OCT visualization. (c) Point cloud obtained by structured light camera (blue) and OCT surface point cloud (sparse sampling). (d) Stitched *en face* OCT image of the phantom surface. (e) Depth-encoded map of the scanned region. (f) and (g) OCT cross-sections at the yellow and red lines in (b), respectively.

the background. Wide-field OCT imaging allowed for comprehensive visualization of the 3D morphology and *ex vivo* structures (Fig. 7(b)). Figure 7(c) shows the color point cloud of the sample surface captured by the structured light camera. The stitched *en face* image provided a broader coverage of surface topography and texture of renal tissue (Fig. 7(d)). This enabled easier identification of surface lesions and abnormalities that may not be easily observable with a single FOV. The depth-encoded map revealed the true surface geometry of the kidney (Fig. 7(e)), while specific locations were analyzed with OCT depth scans (Fig. 7 g, Fig. 7 h with (high resolution image shown in Dataset 1 (Ref. [43]))). Since the experimental subject was *ex vivo* porcine kidney, microscopic structures such as renal tubular lumen cannot be guaranteed to be observed in OCT B-Scan [32]. To preserve the underlying anatomical features, perfusion can be applied for the sample pre-processing.

OCT offers non-invasive imaging of skin, providing clear visualization of microstructures and tissue layers for assessing skin conditions and diseases [6]. The wide-area OCT enables screening and monitoring of large skin areas in patients with dermatological conditions. The reconstructed 3D skin image, as shown in Fig. 8(b), covers an area of 31.5 mm × 15.5 mm. The texture features of the skin and the contour of the scar (indicated by red arrows) are clearly visible (Fig. 8(d)). Figure 8 g and 8 h represent cross-sectional images at the locations indicated by the dashed lines in Fig. 8(b), with the scar marked by a red arrow. Figure 8(f) further zoom in on regions of interest (blue frame in Fig. 8 g) and correspond to scar tissue. The scar tissue exhibits thickened epidermal layers and irregular tissue structures, while the OCT image of normal skin shows uniform tissue structures. The wide-area *en face* and cross-sectional views hold potential in providing valuable information for rapid assessment and identification of the skin's health condition. It is worth mentioning that sample motion will be problematic during the scanning process of several minutes. For *in vivo* biological imaging, such as hand skin, sample fixation was considered in our study. However, slight low-frequency involuntary motion may still occur, resulting in motion artifacts. During the post 3D reconstruction process, registration method

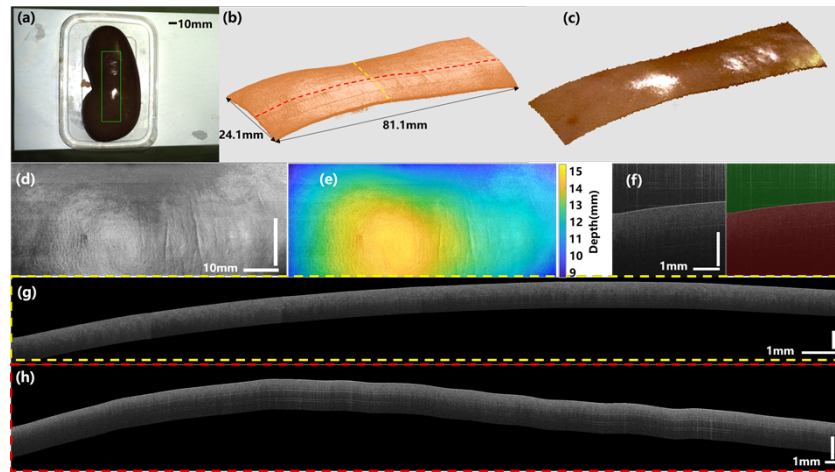


Fig. 7. Results of large area imaging on the isolated kidney. (a) Photographs depicting the isolated kidney and the selection of the region of interest (ROI). (b) 3D OCT visualization. (c) Color point cloud captured by the structured light camera. (d) Stitched *en face* OCT image displaying the surface of the kidney. (e) Depth-encoded map representing the scanned region. (f) B-scan and semantic segmentation results. (g) and (h) OCT cross-sections captured at the yellow and red lines shown in (b), respectively.

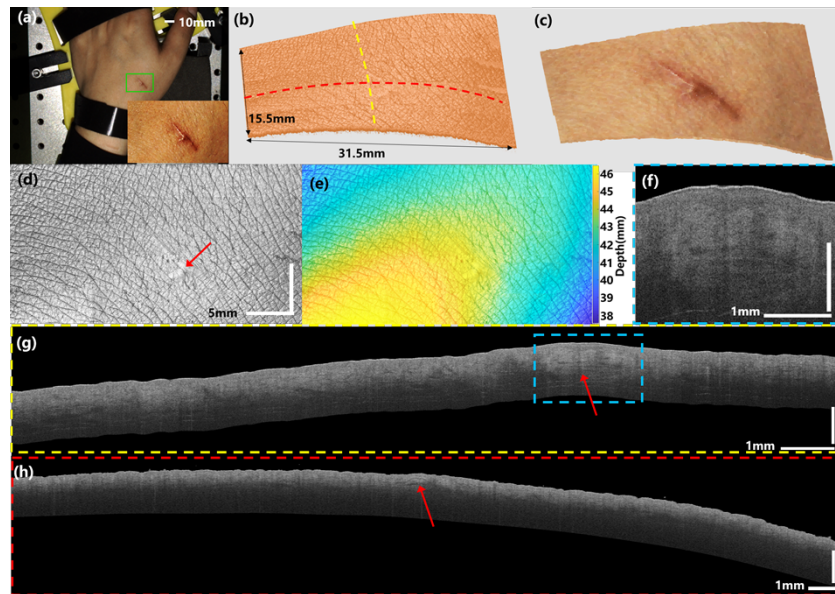


Fig. 8. Large area imaging results of the skin on the back of the hand with scars. (a) The hand was immobilized to minimize tremors during the imaging process. (b) 3D visualization of the hand's skin using OCT imaging. (c) Color point cloud captured by the structured light camera. (d) Stitched *en face* image displaying the surface of the skin, with clearly distinguishable scars indicated by red arrows. (e) Depth-encoded map representing the scanned area. (f) Cross-section of scar tissue. (g), (h) OCT cross-sections captured at the yellow and red lines shown in (b), respectively. High resolution image of (h) shown in Dataset 1 (Ref. [43]).

was applied to correct motion-induced misalignment, thus further reducing the effect of motion artifacts.

3.4. Non-Biomedical Applications

In addition to its extensive application in the medical field, with the advancement of OCT technology, OCT is entering into other domains. OCT enables quantitative analysis of plant morphology and leaf thickness, facilitating the assessment of physiological changes and disease detection in plants [10,39]. We achieved a complete 3D visualization of a leaf using wide-field OCT (Fig. 9(b)). Figure 9(d) reveals the overall shape of the leaf, its edge contour, and the distribution of leaf veins. In practical application scenarios, leaf structures may exhibit extensive changes, such as lesions or defects, which may not be fully captured and analyzed in a single FOV image. In the B-scan images (Fig. 9(f)), the main central vein of the leaf and smaller leaf veins are clearly visible. Figure 9 g, h represent cross-sectional images at the leaf position indicated by the dashed line in Fig. 9(b), allowing for clear differentiation of the leaf (green arrow) and the substrate (blue arrow) in the extensive cross-sectional structures.

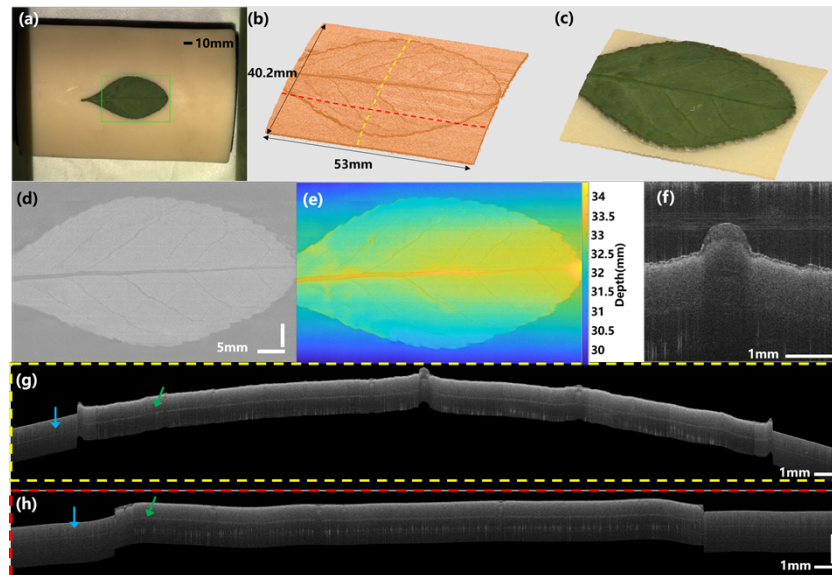


Fig. 9. Large area imaging results of the leaf blade. (a) Photograph of the leaf blade highlighting the scanned area. (b) 3D visualization. (c) Color point cloud captured by the structured light camera. (d) Stitched *en face* image showcasing the surface of the leaf, with distinct visualization of leaf edges and veins. (e) Depth-encoded map representing the scanned area of the leaf. (f) The B-scan image provides a cross-sectional view of the blade structure. (g) and (h) OCT cross-sections captured at the yellow and red lines indicated in (b), respectively. High resolution image of (h) shown in Dataset 1 (Ref. [43]).

In addition, OCT has been applied to evaluate fruit quality and non-destructively detect internal defects and diseases [11]. The use of a robotic arm-assisted wide-field OCT enables rapid scanning of a large surface area of the fruit, providing comprehensive detection information. We obtained a wide-field 3D visualization of a scarred citrus peel (Fig. 10(b)). Extensive *en face*

imaging (Fig. 10(d)) revealed detailed features including oil glands and scar tissue on the fruit's surface, while depth-encoded maps reflected the true topography of the citrus peel. Examples of OCT depth scans are shown in Fig. 10 g, h to display the overall depth profile. Further magnified images of specific regions (Fig. 10(f)), indicated by the blue boxes in Fig. 10 g, demonstrated clear observation of scar tissue.

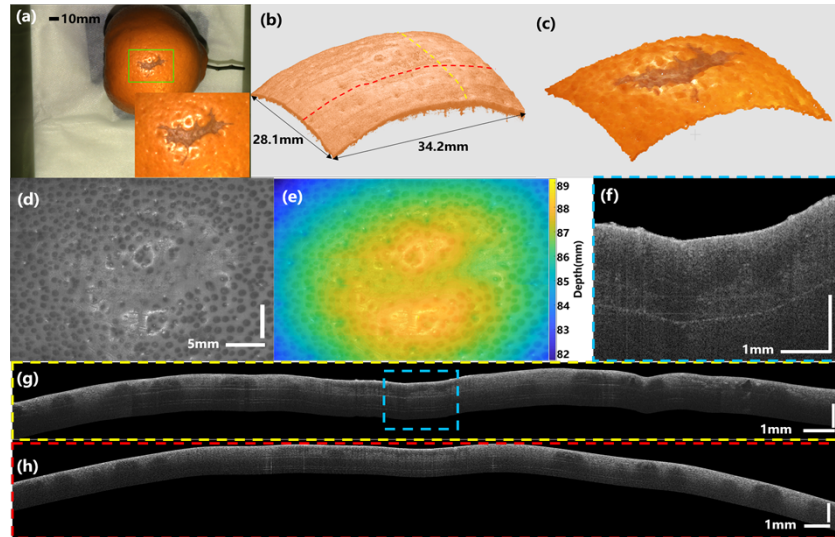


Fig. 10. Large area imaging results of citrus. (a) Photograph depicting the citrus sample and the corresponding scanned area. (b) 3D visualization. (c) Color point cloud captured by the structured light camera. (d) Stitched *en face* image showcasing the surface of the citrus, with clear visualization and distinction of the scars. (e) Depth-encoded map representing the scanned area of the citrus. (f) B-scan images depicting cross-sections of scar tissue. (g) and (h) OCT cross-sections captured at the yellow and red lines indicated in (b), respectively. High resolution image of (h) shown in [Dataset 1](#) (Ref. [43]).

OCT technology also finds applications in materials science and cultural heritage preservation, enabling high-precision surface defect and internal crack detection in ceramic materials [41,42]. It allows for detailed analysis of the structure and thickness of glaze and body. Wide-field OCT enables non-destructive imaging and inspection of large-sized ceramics. We performed a comprehensive 3D reconstruction of the sidewall of a ceramic cup (Fig. 11(b)). Figure 11(f) demonstrates the ability of OCT to accurately differentiate between different structures and thicknesses of the glaze and ceramic body. Researchers can extract microstructural information, detect secondary repairs, and identify internal defects from large-scale OCT cross-sectional images (Fig. 11 (g), (h)). *En face* images (Fig. 11(d)) provide insights into crystalline particles within the glaze and surface defects on the ceramic body. The bright spot indicated by the red arrow represents a strong reflection on the surface.

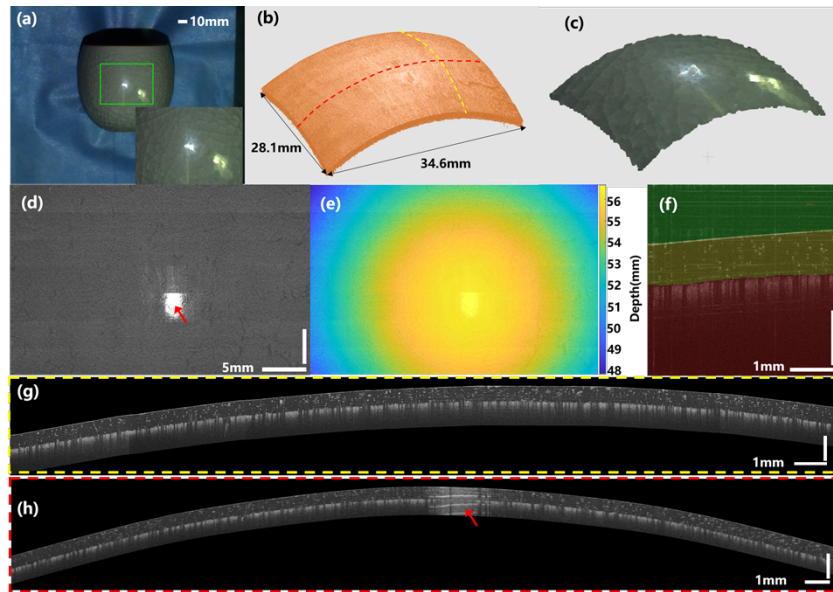


Fig. 11. Large area imaging results of ceramics. (a) Photograph illustrating the ceramic sample and the corresponding scanned area. (b) 3D visualization. (c) Color point cloud captured by the structured light camera. (d) Stitched *en face* image of the ceramic surface, showcasing areas of high surface reflection and brightness. (e) Depth-coded map representing the scanned area of the ceramics, facilitating depth visualization. (f) Semantic segmentation results with the background portion highlighted in green, the enamel portion highlighted in yellow, and the base highlighted in red. (g) and (h) OCT cross-sections captured at the yellow and red lines indicated in (b), respectively, allowing detailed analysis at specific locations within the ceramics. High resolution image of (h) shown in [Dataset 1](#) (Ref. [43]).

4. Conclusion and discussion

In summary, we have proposed an automatic robot-assisted large-area OCT system based on structured light camera guidance. Compared to manual path planning and RGB depth camera and OCT image hybrid guidance that requires landing-motion and continuous scan path modification via closed-loop OCT feedback [32], the system enables efficient and accurate large-area OCT scanning and generates 3D reconstructions of the sample with wide FOV and high resolution. We have demonstrated the accuracy and versatility of the proposed system through position error tests and examples of 3D macroscopic imaging of different samples. The cost of the wide FOV and high resolution is the imaging time, which can be reduced by using a faster OCT system or lowering the sampling rate. Furthermore, the combination of the system with AI semantic algorithms for real-time image segmentation processing can better display the reconstructed image. Future work could involve using more powerful segmentation algorithms to extract diagnostically significant information and plan of free scan path instead of raster scan path.

The proposed system provides a good solution to overcome the limitations of traditional OCT systems and may open up new opportunities for large-area OCT imaging in various fields, including biomedical and non-biomedical applications, with broad prospects for future development and application.

Funding. National Key Research and Development Program of China (2022YFB4702902); National Natural Science Foundation of China (62275023); Beijing Municipal Natural Science Foundation (4232077); Overseas Expertise Introduction Project for Discipline Innovation (B18005).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in [Dataset 1](#), Ref. [43].

References

1. B. E. Bouma, J. F. de Boer, D. Huang, I. K. Jang, T. Yonetsu, C. L. Leggett, R. Leitgeb, D. D. Sampson, M. Suter, B. Vakoc, M. Villiger, and M. Wojtkowski, "Optical coherence tomography," *Nat. Rev. Methods Primers* **2**(1), 79 (2022).
2. S. Zheng, Y. Bai, Z. Xu, P. Liu, and G. Ni, "Optical coherence tomography for three-dimensional imaging in the biomedical field: a review," *Front. Phys.* **9**, 744346 (2021).
3. M. Li, S. Landahl, A. R. East, P. Verboven, and L. A. Terry, "Optical coherence tomography—A review of the opportunities and challenges for postharvest quality evaluation," *Postharvest Biol. Technol.* **150**, 9–18 (2019).
4. W. Drexler and J. G. Fujimoto, "Optical coherence tomography: technology and applications," in *Springer Science & Business Media* (Springer, 2015), Chap. 5, p. 169.
5. C.P. Herbot Jr., M. Takeuchi, I. Papasavvas, I. Tugal-Tutkun, A. Hedayatfar, Y. Usui, P.C. Ozdal, and C.A. Urzua, "Optical coherence tomography angiography (OCT-A) in uveitis: a literature review and a reassessment of its real role," *Diagnostics* **13**(4), 601 (2023).
6. J. Olsen, J. Holmes, and G. B. E. Jemec, "Advances in optical coherence tomography in dermatology—a review," *J. Biomed. Opt.* **23**(04), 1 (2018).
7. J. Wang, Y. Xu, and S. A. Boppart, "Review of optical coherence tomography in oncology," *J. Biomed. Opt.* **22**(12), 1 (2017).
8. F. N. Mohamad Saberi, P. Sukumaran, N. M. Ung, and Y. M. Liew, "Assessment of demineralized tooth lesions using optical coherence tomography and other state-of-the-art technologies: a review," *Biomed. Eng. Online* **21**(1), 83 (2022).
9. C. S. Cheung, M. Spring, and H. Liang, "Ultra-high resolution fourier domain optical coherence tomography for old master paintings," *Opt. Express* **23**(8), 10145–10157 (2015).
10. J. de Wit, S. Tonn, G. Ackerveken, and J. Kalkman, "Quantification of plant morphology and leaf thickness with optical coherence tomography," *Appl. Opt.* **59**(33), 10304–10311 (2020).
11. R. E. Wijesinghe, S.-Y. Lee, N. K. Ravichandran, M. F. Shirazi, P. Kim, H. Jung, M. Jeon, and J. Kim, "Biophotonic approach for the characterization of initial bitter-rot progression on apple specimens using optical coherence tomography assessments," *Sci. Rep.* **8**(1), 15816 (2018).
12. M. Everett, S. Magazzini, T. Schmoll, and M. Kempe, "Optical coherence tomography: From technology to applications in ophthalmology," *Transl. Biophotonics* **3**(1), e202000012 (2021).
13. F. Zheng, X. Deng, Q. Zhang, J. He, P. Ye, S. Liu, P. Li, J. Zhou, and X. Fang, "Advances in swept-source optical coherence tomography and optical coherence tomography angiography," *Adv. Ophthalmol. Pract. Res.* **3**(2), 67–79 (2023).
14. M. Ripa, L. Motta, T. Florit, J.-Y. Sahyoun, V. Matello, and B. Parolini, "The role of widefield and ultra widefield optical coherence tomography in the diagnosis and management of vitreoretinal diseases," *Diagnostics* **12**(9), 2247 (2022).
15. J. Xu, S. Song, S. Men, and R. K. Wang, "Long ranging swept-source optical coherence tomography-based angiography outperforms its spectral-domain counterpart in imaging human skin microcirculations," *J. Biomed. Opt.* **22**(11), 1–11 (2017).
16. J. Lefebvre, A. Castonguay, P. Pouliot, M. Descoteaux, and F. Lesage, "Whole mouse brain imaging using optical coherence tomography: reconstruction, normalization, segmentation, and comparison with diffusion MRI," *Neurophotonics* **4**(4), 041501 (2017).
17. B. He, Y. Zhang, Z. Meng, Z. He, Z. Chen, Z. Yin, Z. Hu, Y. Shi, C. Wang, X. Zhang, N. Zhang, G. Wang, and P. Xue, "Whole brain micro-vascular imaging using robot assisted optical coherence tomography angiography," *IEEE J. Sel. Top. Quantum Electron.* **29**(4: Biophotonics), 1–9 (2023).
18. T. Callewaert, J. Guo, G. Harteveld, A. Vandivere, E. Eisemann, J. Dik, and J. Kalkman, "Multi-scale optical coherence tomography imaging and visualization of Vermeer's Girl with a Pearl Earring," *Opt. Express* **28**(18), 26239–26256 (2020).
19. Z. Wang, B. Potsaid, L. Chen, C. Doerr, H.-C. Lee, T. Nielson, V. Jayaraman, A. E. Cable, E. Swanson, and J. G. Fujimoto, "Cubic meter volume optical coherence tomography," *Optica* **3**(12), 1496–1503 (2016).
20. J. Xu, S. Song, W. Wei, and R. K. Wang, "Wide field and highly sensitive angiography based on optical coherence tomography with a kinetic swept source," *Biomed. Opt. Express* **8**(1), 420–435 (2017).
21. F. Schwarzhans, S. Desissaire, S. Steiner, M. Pircher, C. K. Hitzenberger, H. Resch, C. Vass, and G. Fischer, "Generating large field of view *en face* projection images from intra-acquisition motion compensated volumetric optical coherence tomography data," *Biomed. Opt. Express* **11**(12), 6881–6904 (2020).
22. M. H. Laves, L. A. Kahrs, and T. Ortmaier, "Volumetric 3D stitching of optical coherence tomography volumes," *Curr. Dir. Biomed. Eng.* **4**(1), 327–330 (2018).
23. Y. Ji, K. Zhou, S. H. Ibbotson, R. K. Wang, C. Li, and Z. Huang, "A novel automatic 3D stitching algorithm for optical coherence tomography angiography and its application in dermatology," *J. Biophotonics* **14**(11), e202100152 (2021).
24. J. B. Eom, J. Ahn, J. Eom, and A. Park, "Wide field of view optical coherence tomography for structural and functional diagnoses in dentistry," *J. Biomed. Opt.* **23**(07), 1 (2018).

25. J. Walther, J. Golde, M. Albrecht, B. Quirk, L. Scolaro, R. W. Kirk, Y. Gruda, C. Schnabel, F. Tetschke, K. Joehrens, D. Haim, M. Buckova, J. Li, and R. A. McLaughlin, "A handheld fiber-optic probe to enable optical coherence tomography of oral soft tissue," *IEEE Trans. Biomed. Eng.* **69**(7), 2276–2282 (2022).
26. C. Viehland, X. Chen, D. Tran-Viet, M. Jackson-Atogi, P. Ortiz, G. Waterman, L. Vajzovic, C. A. Toth, and J. A. Izatt, "Ergonomic handheld OCT angiography probe optimized for pediatric and supine imaging," *Biomed. Opt. Express* **10**(5), 2623–2638 (2019).
27. J. Yang, L. Liu, J. Campbell, D. Huang, and G. Liu, "Handheld optical coherence tomography angiography," *Biomed. Opt. Express* **8**(4), 2287–2300 (2017).
28. Y. Huang, X. Li, J. Liu, Z. Qiao, J. Chen, and Q. Hao, "Robotic-arm-assisted flexible large field-of-view optical coherence tomography," *Biomed. Opt. Express* **12**(7), 4596–4609 (2021).
29. M. Göb, S. Lotz, L. Ha-Wissel, S. Burhan, S. Böttger, F. Ernst, J. Hundt, and R. Huber, "Large area robotically assisted optical coherence tomography (LARA-OCT) for skin imaging with MHz-OCT surface tracking," *Proc. SPIE* **12367**, 29 (2023).
30. L. Zhang, M. Ye, P. Giataganas, M. Hughes, and G. -Z. Yang, "Autonomous scanning for endomicroscopic mosaicing and 3D fusion," *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 3587–3593 (2017).
31. P. Ortiz, M. Draelos, A. Narawane, R. P. McNabb, A. N. Kuo, and J. A. Izatt, "Robotically-Aligned Optical Coherence Tomography with Gaze Tracking for Live Image Montaging of the Retina," *2022 International Conference on Robotics and Automation (ICRA)*, Philadelphia, PA, USA, 3783–3789 (2022).
32. X. Ma, M. Moradi, H. Mustafa, M. Hunter, Y. Chen, and H. K. Zhang, "Feasibility of robotic-assisted optical coherence tomography with extended scanning area for pre-transplant kidney monitoring," *Proc. SPIE* **11948**, 24 (2022).
33. Y. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX = XB$," *IEEE Trans. Robot. Automat.* **5**(1), 16–29 (1989).
34. R. Tsai and R. Lenz, "A new technique for fully autonomous and efficient 3d robotics hand/eye calibration," *IEEE Trans. Robot. Automat.* **5**(3), 345–358 (1989).
35. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.* **9**(1), 62–66 (1979).
36. R. Grompone von Gioi and G. Randall, "A sub-pixel edge detector: an implementation of the canny/levernay algorithm," *Image Process. Line* **7**, 347–372 (2017).
37. C. Yu, C. Gao, J. Wang, G. Yu, C. Shen, and N. Sang, "BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation," *Int. J. Comput. Vis.* **129**(11), 3051–3068 (2021).
38. B. Konkel, C. Lavin, T. Wu, E. Anderson, A. Iwamoto, H. Rashid, B. Gaitian, J. Boone, M. Cooper, P. Abrams, A. Gilbert, Q. Tang, M. Levi, J. G. Fujimoto, P. Andrews, and Y. Chen, "Fully automated analysis of OCT imaging of human kidneys for prediction of post-transplant function," *Biomed. Opt. Express* **10**(4), 1794–1821 (2019).
39. R. E. Wijesinghe, S.-Y. Lee, P. Kim, H.-Y. Jung, M. Jeon, and J. Kim, "Optical Inspection and Morphological Analysis of *Diospyros kaki* Plant Leaves for the Detection of Circular Leaf Spot Disease," *Sensors* **16**(8), 1282 (2016).
40. R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," *2011 IEEE International Conference on Robotics and Automation*, Shanghai, China, 1–4 (2011).
41. R. Su, M. Kirillin, E. W. Chang, E. Sergeeva, S. H. Yun, and L. Mattsson, "Perspectives of mid-infrared optical coherence tomography for inspection and micrometrology of industrial ceramics," *Opt. Express* **22**(13), 15804–15819 (2014).
42. M.-L. Yang, J. I. Katz, J. Barton, W.-L. Lai, and J.-H. Jean, "Using Optical Coherence Tomography to Examine Additives in Chinese Song Jun Glaze," *Archaeometry* **57**(5), 837–855 (2015).
43. X. Li, "Original high resolution," figshare (2023), <https://doi.org/10.6084/m9.figshare.23643033>.